# Part A

1. Try training the neural network by clicking the run button (top left). How quickly does it find a solution for each of the four sample classification tasks?   
   *Note: the times below are given at when the* ***test\_loss*** *and* ***training\_loss*** *are as close to 0 as possible, but no more than 0.02. 70/30 split with batch\_size of 10.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Task | | Time 1 | Time 2 | Time 3 |
| Circle |  | 932 epochs | 241 epochs | 401 epochs |
| Exclusive OR |  | 1210 epochs | 1200 epochs, test\_loss ~0.007 | 980 epochs |
| Gaussian |  | 141 epochs | 157 epochs | 159 epochs |
| Spiral |  | 3k epoch stuck on optima | No convergence after 6k epoch | No convergence after 6k epoch |

1. Try replacing the tanh activation function with the reLU function and train the network again. What can you say about the output (shape and time taken to get a good classification)?  
   Batch size: 15 train:test ratio: 70%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Task | | Time 1 | Time 2 | Time 3 |
| Circle | A picture containing text  Description automatically generated | 480 epochs | 377 epochs | 416 epochs |
| Exclusive OR | Qr code  Description automatically generated | 900 epochs | 565 epochs | 182 epochs |
| Gaussian | A picture containing text, light, image, dark  Description automatically generated | 211 epochs | 71 epochs | 196 epochs |
| Spiral | A picture containing text, clipart  Description automatically generated | >5k epochs; can’t learn | 3.5k epochs; can’t learn | 3-4k oscillation: |

1. Modify the network architecture to have just one hidden layer with three neurons. Train it multiple times (to reset the network weights, click the reset button). What can you comment about this training vs. your answer in (a.)?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Task | | Time 1 | Time 2 | Time 3 |
| Circle | A picture containing text  Description automatically generated | 631 epochs | 728 epochs | 664 epochs |
| Exclusive OR | Qr code  Description automatically generated | 687 epochs | 700 Epochs | 642 epochs |
| Gaussian | A picture containing text, light, image, dark  Description automatically generated | 29 epochs | 26 epochs | 26 epochs |
| Spiral | A picture containing text, clipart  Description automatically generated | >2k epochs | >1k epochs | >2k epochs |

Of note is that the configuration of the 3 neurons in the hidden layer will vastly change the end result. Exclusive OR’s Time 2, all of the neurons were covering primarily the right half of the domain.

1. Remove one neuron to keep just two. Retrain the network. What can you comment about this training vs. your answer in (a.) and (c.)?

Running each of the classification tasks with a network featuring just two networks in its sole hidden layer reveals how with a small count of neurons, we may have overlapping neurons that “double up” preventing convergence. This problem goes away quickly as the number of hidden layers and neurons increase.   
  
Of all the tasks, the Gaussian distribution was the only classification that managed to remain successful. In fact, it reaches 0.001 train/test loss in under 15 epochs; the fastest we’ve seen yet.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Circle | XOR | Gaussian | Spiral |
| Time to converge / Shape |  |  | 14 epochs! |  |

1. Set the number of neurons to eight and train the network several times.

Using the max number of neurons allows the algorithm to all but the spiral within 100 epochs. The spiral takes significantly longer to train since its such a complicated set.

A screenshot of a graph

Description automatically generated with medium confidence